

Identifiability of an X-rank decomposition of polynomial maps*

Pierre Comon, Yang Qi, and Konstantin Usevich

pierre.comon@gipsa-lab.fr, yangqi@galton.uchicago.edu,

konstantin.usevich@gipsa-lab.fr

Abstract. In this paper, we study a polynomial decomposition model that arises in problems of system identification, signal processing and machine learning. We show that this decomposition is a special case of the X-rank decomposition — a powerful novel concept in algebraic geometry that generalizes the tensor CP decomposition. We prove new results on generic/maximal rank and on identifiability of a particular polynomial decomposition model. In the paper, we try to make results and basic tools accessible for general audience (assuming no knowledge of algebraic geometry or its prerequisites).

Key words. X-rank, identifiability, polynomial decomposition, Waring decomposition, generic rank

AMS subject classifications. 12E05; 14M12; 15A21; 15A69

1. Introduction: polynomial decompositions.

1.1. Notation. We use boldface letters ($\mathbf{a}, \mathbf{b}, \dots$) for vectors, and boldface capital letters ($\mathbf{A}, \mathbf{B}, \dots$) for matrices. Given an m -dimensional vector space A over a field \mathbb{K} , fix a basis for A , then a vector $\mathbf{a} \in A$ can be identified with an $m \times 1$ matrix, i.e., $\mathbf{a} = [a_1 \ \cdots \ a_m]^\top$, where \cdot^\top denotes the transpose. Thus, $\mathbf{a}^\top \mathbf{b}$ stands for the matrix multiplication¹ $\mathbf{a}^\top \mathbf{b} = a_1 b_1 + \cdots + a_m b_m$. By Π_m^d we denote the space of multivariate polynomials in m variables of total degree $\leq d$, and we write an element of Π_m^d in the form $f(\mathbf{u})$, where $\mathbf{u} = [u_1 \ \cdots \ u_m]^\top$.

Standardly, we use \times for Cartesian product of sets, and a shorthand notation $A^{\times d} = A \times \cdots \times A$. We use $A \oplus B$ for the direct sum² of vector spaces, and \otimes for the tensor product. By $S^d(V)$ or $S^d V$ we denote the space of d -th order symmetric tensors on an m -dimensional vector space V (i.e., $m \times \cdots \times m$ symmetric tensors). In $S^d V$, \mathbf{v}^d means $\mathbf{v} \otimes \cdots \otimes \mathbf{v}$.

1.2. Model and examples. Let \mathbb{K} be \mathbb{R} or \mathbb{C} . Consider a multivariate polynomial map $\mathbf{f} : \mathbb{K}^m \rightarrow \mathbb{K}^n$, i.e., a vector $\mathbf{f}(\mathbf{u}) = [f_1(\mathbf{u}) \ \cdots \ f_n(\mathbf{u})]^\top \in (\Pi_m^d)^{\times n}$ of multivariate polynomials of total degree $\leq d$ in m variables, (i.e., each $f_i \in \Pi_m^d$). Without loss of generality, in this paper, we assume that $f_k(\mathbf{0}) = \mathbf{0}$ (i.e., the constant part of \mathbf{f} is zero).

Following [20], we say that \mathbf{f} has a *decoupled representation*, if it can be expressed as

$$\mathbf{f}(\mathbf{u}) = \mathbf{w}_1 g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + \mathbf{w}_r g_r(\mathbf{v}_r^\top \mathbf{u}), \quad (1.1)$$

where $\mathbf{v}_k \in \mathbb{K}^m$, $\mathbf{w}_k \in \mathbb{K}^n$, and where $g_k(t) = c_{1,k}t + \cdots + c_{d,k}t^d$ are univariate polynomials over \mathbb{K} . The problem is often to find a decoupled representation (1.1) with r minimum.

*This work is supported by the ERC project “DECODA” no.320594, in the frame of the European program FP7/2007-2013.

¹Note that this is not the inner product in the case $\mathbb{K} = \mathbb{C}$

²i.e. the Cartesian product $A \times B$ equipped with the vector space structure

Example 1 ($d = 1$). In this case, \mathbf{f} is a linear map, i.e. $\mathbf{f}(\mathbf{u}) = \mathbf{F} \cdot \mathbf{u}$ with $\mathbf{F} \in \mathbb{K}^{n \times m}$. Without loss of generality we can assume $g_k(t) = t$, and (1.1) becomes a low-rank factorization³

$$\mathbf{F} = \mathbf{w}_1 \mathbf{v}_1^\top + \cdots + \mathbf{w}_r \mathbf{v}_r^\top.$$

The next special case is one of the key examples in this paper.

Example 2 ($n = 1$). In this case \mathbf{f} is a single polynomial $\mathbf{f}(\mathbf{u}) = f(\mathbf{u})$, and (1.1) becomes

$$f(\mathbf{u}) = g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + g_r(\mathbf{v}_r^\top \mathbf{u}), \quad (1.2)$$

since we can assume that $\mathbf{w}_k = [1]$. An example of (1.2) is shown in fig. 1.1d. The decomposition (1.2)

- is known as sum of *ridge functions* or plane waves [29, 31] in approximation theory;
- corresponds to ridge polynomial neural networks [36] (RPNs) in machine learning;
- appears in blind source separation problems in signal processing [17].

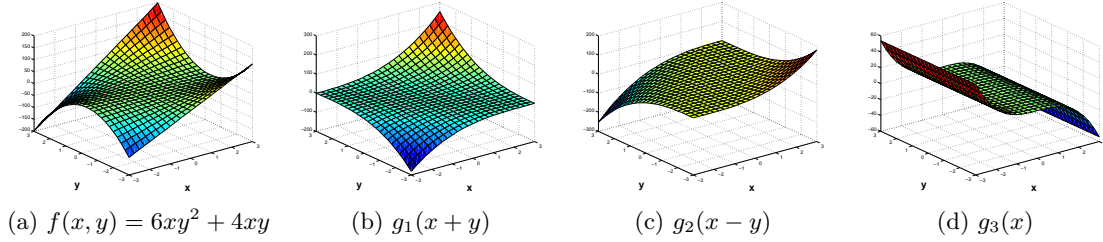


Figure 1.1. $f(x, y) = g_1(x + y) + g_2(x - y) + g_3(x)$, $g_1(t) = t^3 + t^2$, $g_2(t) = t^3 - t^2$, $g_3(t) = -2t^3$.

Next, the homogeneous versions of eq. (1.1) and eq. (1.2) are well-known in algebraic geometry.

Example 3 ($n = 1$, f — homogeneous). If f is homogeneous of degree d , then $g_k(t)$ should be also homogeneous, i.e. $g_k(t) = c_k t^d$. Hence, the decomposition (1.1) becomes

$$f(\mathbf{u}) = c_1 \cdot (\mathbf{v}_1^\top \mathbf{u})^d + \cdots + c_r \cdot (\mathbf{v}_r^\top \mathbf{u})^d. \quad (1.3)$$

The decomposition (1.3) is known as *Waring decomposition*, and was subject to numerous studies in the literature [26, 1]. Via the correspondence between homogeneous polynomials and symmetric tensors (see appendix A.1), (1.3) becomes the *symmetric tensor decomposition*

$$f = c_1 \mathbf{v}_1^d + \cdots + c_r \mathbf{v}_r^d, \quad (1.4)$$

where $f \in S^d V$ is the symmetric tensor corresponding to the polynomial in $f(\mathbf{u})$. For homogeneous case, the general decomposition (for $n > 1$) was also already considered.

Example 4 ($n > 1$, f — homogeneous). As in example 3, (1.1) can be rewritten as

$$f(\mathbf{u}) = \mathbf{w}_1 \cdot (\mathbf{v}_1^\top \mathbf{u})^d + \cdots + \mathbf{w}_r \cdot (\mathbf{v}_r^\top \mathbf{u})^d. \quad (1.5)$$

³example 1 shows that (1.1) can be interpreted as a “low-rank factorization” of a nonlinear map.

The decomposition (1.5) is exactly the simultaneous Waring decomposition of homogeneous polynomials $f_1(\mathbf{u}), \dots, f_n(\mathbf{u})$ (equivalently, CP decomposition of a partially symmetric tensor).

Example 5 (the general case, $n > 1$, \mathbf{f} — non-homogeneous). As summarized in [20], the general decomposition (1.1) appears in the field of nonlinear system identification [35, 24]. A common problem in identification (parameter estimation) for several challenging nonlinear block-structured systems (parallel Wiener-Hammerstein [35] and nonlinear feedback [40] models) is to decompose a nonlinear function (represented by a polynomial) in the form (1.1).

Remark 1. In the system identification literature ([20]), the decomposition (1.1) is often written in a compact form

$$\mathbf{f}(\mathbf{u}) = \mathbf{W}\mathbf{g}(\mathbf{V}^\top \mathbf{u}),$$

where $\mathbf{V} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_r] \in \mathbb{K}^{m \times r}$, $\mathbf{W} = [\mathbf{w}_1 \ \dots \ \mathbf{w}_r] \in \mathbb{K}^{n \times r}$ and $\mathbf{g} : \mathbb{K}^r \rightarrow \mathbb{K}^r$ defined as $\mathbf{g}(t_1, \dots, t_r) = [g_1(t_1) \ \dots \ g_r(t_r)]^\top$. Also, a block-diagram for decomposition (1.1) (given in fig. 1.2) is often used, where the “input” variables \mathbf{u} are transformed by a linear transformation, followed by component-wise nonlinear transformations. The “outputs” are obtained by linear combinations of the results of the nonlinear transformation.

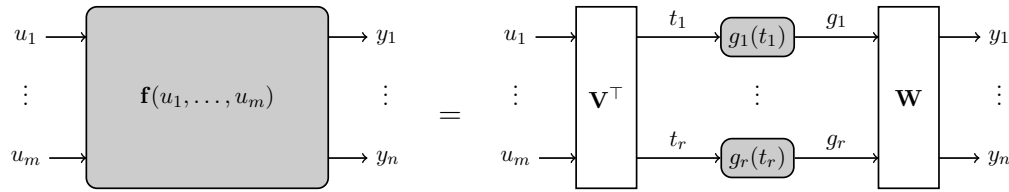


Figure 1.2. Representation of a polynomial decomposition.

1.3. Goals and previous works. When using model (1.1), a few natural theoretical questions arise that are important to understand the limits of the applicability of the model.

1. When is the model *identifiable*? (i.e., when is the decomposition (1.1) unique?).
2. What is the upper bound on r in (1.1) needed to represent any polynomial?
3. What is the typical (for a “random” \mathbf{f}) behavior of r in the shortest decomposition?

As for the special (homogeneous) cases of decomposition (1.1) (Examples 1,3,4), all the three cases were a subject of rapid development in the last two decades, and many results are available. In this paper, we address the non-homogeneous case (Examples 2 and 5), where very few results are available (listed below).

Bounds on r and typical behavior. This question was considered only for $n = 1$, in the papers [33, 34, 5]. The best result shows that any $f \in \Pi_m^d$ can be decomposed as (1.2) whenever

$$r \leq \binom{m+d-2}{d-1}, \quad (1.6)$$

where the bound⁴ (1.6) is valid for \mathbb{R} , \mathbb{C} and for certain finite fields. The typical behavior of r in the shortest decomposition is known only for the case $m = 2$ and $n = 1$ [33] (the case of bivariate polynomials).

Uniqueness. The uniqueness in representations (1.1) was almost not studied. The authors of [20] suggested to construct a structured tensor from the coefficients of polynomials. Based on a Kruskal-type condition for unstructured tensors, they propose a bound for generic uniqueness that depends on r, m, d . This bound is, however, applicable only to unstructured tensors, and not to the decomposition (1.1), as we argue in remark 14.

Remark 2. *A common idea (suggested to us by one of the reviewers) is that decomposition (1.2) can be brought to the form (1.3), and hence Waring decomposition can be applied (the same argument can be applied to bring (1.1) to the form (1.5)). However, homogenization can increase the number of terms, and does not give a good answer to our questions.*

For example, the homogenization of $f(x, y)$ in fig. 1.1d is the trivariate polynomial

$$6xy^2 + 4xy \xrightarrow{\text{homogenization}} 6xy^2 + 4xyz = xy(6y + z).$$

But it is known [8] that this homogeneous polynomial does not have a Waring decomposition eq. (1.3) with less than 4 terms (compare with 3 terms in fig. 1.1d). The reason for that is that the polynomials g_1, g_2, g_3 do not correspond to powers of linear forms for the homogenized polynomial. In fact, homogenization restricts the form of polynomials g_k .

1.4. Contribution and structure of this paper. In this paper, we show that the decomposition (1.1) can be viewed as a special case of X -rank decomposition. The notion of X -rank (or rank with respect to a variety \tilde{X}) is a powerful concept developed in the field of algebraic geometry that generalizes matrix rank, tensor rank, symmetric tensor rank and other notions of rank. The questions raised in section 1.3 can be addressed in the framework of X -rank and correspond to finding maximal, typical, generic ranks and to checking r -identifiability (generic uniqueness). In particular, we:

1. Obtain results on identifiability and partial identifiability of (1.1).
2. Determine the value of generic rank for some special cases of $n = 1$.
3. Obtain a new bound on r_{\max} (for $\mathbb{K} = \mathbb{R}$ or \mathbb{C}) that is better than (1.6).

Although in this paper we do not develop decomposition algorithms (see [20], [40], [39] for available algorithms), we believe that the ideas may lead to new or improved algorithms.

In section 2, we introduce the concept of X -rank decompositions and make a review of recent results. We prefer a very simplistic exposition and hope that section 2 may serve as an entry point to the literature on X -rank for a wider audience, including applied mathematicians and engineers. In section 3, we recall the definition and known results on generic uniqueness (identifiability), and prove equivalence of different definitions appearing in the literature. In section 4, we introduce Veronese scrolls, show that decompositions (1.1) and (1.2) are related to X -rank decompositions for Veronese scrolls, and give defining equations for this variety. Section 5 contains the main results of the paper, including identifiability of Veronese scrolls and polynomial decompositions, dimensions of secant varieties, and results on generic ranks.

⁴Bound (1.6) is better than a naive bound $\binom{m+d-1}{d}$ (number of monomials in the highest degree part of f).

2. X-rank decompositions. The concept of X -rank (or rank with respect to a variety) was probably first proposed in [41], and popularized in [6, 27]. In this section we give key definitions and basic results, in a simplified form. In particular, we avoid the use of projective varieties whenever possible.

2.1. X-rank: definitions. Consider an N -dimensional vector space⁵ A over \mathbb{K} , where \mathbb{K} is \mathbb{R} or \mathbb{C} . Assume that a subset $\hat{X} \subset A$ is fixed that satisfies the following conditions.

Assumption 1. \hat{X} is scale-invariant, i.e. $\mathbf{v} \in \hat{X}$ and $\alpha \in \mathbb{K}$ implies $\alpha \mathbf{v} \in \hat{X}$.

Assumption 2. \hat{X} is non-degenerate, i.e. it is not contained in any hyperplane of A .

Assumption 3. \hat{X} is an algebraic variety, i.e. the zero set of a system of polynomial equations (see also appendix A.2).

Definition 2.1. Given a subset $\hat{X} \subset A$, the X -rank of any vector $\mathbf{v} \in A$ is defined as the smallest number of rank-one elements, such that \mathbf{v} can be represented as their sum:

$$\text{rank}_X(\mathbf{v}) = \min r : \mathbf{v} = \mathbf{x}_1 + \cdots + \mathbf{x}_r, \quad \mathbf{x}_k \in \hat{X}. \quad (2.1)$$

Such a decomposition with the minimal possible number of terms is called the X -rank decomposition. (The rank of $\mathbf{0} \in A$, by convention, is zero.)

Assumption 1 guarantees that the X -rank is compatible with linear operations, whereas Assumption 2 ensures that any vector has an X -rank decomposition and that the X -rank does not exceed N . The Assumption 3 allows for an algebraic analysis of X -rank decompositions.

The X -rank decomposition can be illustrated in fig. 2.1. It is also similar in spirit to sparse (atomic) decompositions, that appeared recently in other branches of applied mathematics [10].

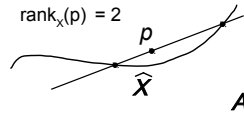


Figure 2.1. Vector p can be decomposed into the sum of 2 elements of the variety \hat{X} .

Remark 3. In fact, Assumptions 1 and 3 imply that \hat{X} is an affine cone of a projective algebraic variety⁶ $X \subset \mathbb{P}A$. The projective variety X is the usual starting point in the definition of X -rank, see [41, 6, 27]. In this paper, however, we prefer to work and give definitions in terms of the affine variety \hat{X} , which simplifies some expressions (as we will show later). One only has to bear in mind that $\dim X = \dim \hat{X} - 1$. To avoid pathological phenomena and also for convenience of using algebraic geometry, the following assumption is often imposed.

Assumption 4. \hat{X} is an irreducible variety (see appendix A.2). Finally, for real varieties, the following assumption is often added, to avoid unexpected phenomena and make use of the powerful tools from complex algebraic geometry.

Assumption 5. A real variety $\hat{X}_{\mathbb{R}}$ has a smooth point in its complexification⁷ $\hat{X}_{\mathbb{R}} \otimes \mathbb{C}$.

⁵For simplicity, one can think that $A = \mathbb{K}^N$.

⁶where $\mathbb{P}A$ is the projective space.

⁷See appendix A.2 for definition of complexification.

2.2. Examples. The basic examples, considered in example 1, example 3 and example 4 fit in the framework of X -rank, and are explained in table 2.1

Table 2.1
Varieties and X -ranks

Ambient space (A)	$\dim(A)$	variety \widehat{X}	$\dim(\widehat{X})$
$\mathbb{K}^{I_1} \otimes \cdots \otimes \mathbb{K}^{I_d}$ tensor	$I_1 \cdots I_d$	$\text{Seg}(\mathbb{K}^{I_1} \times \cdots \times \mathbb{K}^{I_d}) = \{\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_d\}$ Segre variety	$\sum_{k=1}^d I_k - d + 1$
$S^d(\mathbb{K}^m)$ symmetric tensor	$\binom{m+d-1}{d}$	$\nu_d(\mathbb{K}^m) = \{c\mathbf{a}^d\}$ Veronese variety	m
$(S^d(\mathbb{K}^m))^{\times n}$ several symmetric tensors	$n \binom{m+d-1}{d}$	$\text{Seg}(\mathbb{K}^n \times \nu_d(\mathbb{K}^m)) = \{\mathbf{w} \otimes \mathbf{a}^d\}$ Segre-Veronese variety	$m + n - 1$

All the examples in table 2.1 satisfy Assumptions 1 to 5.

The dimension of the variety of rank-one elements \widehat{X} reflects the number of degrees of freedom in the parameterization of \widehat{X} . Take, for instance, the case of non-symmetric tensors (1-st row in table 2.1). It is parameterized by $I_1 + \cdots + I_d$ parameters, but there are $d - 1$ redundancies since any element of \widehat{X} has many representations in the form $\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_d$, due to exchange of scaling. The other examples in table 2.1 follow the same pattern: the dimension of \widehat{X} is equal to the number of parameters minus the number of “dependencies”.

2.3. Maximal, typical ranks and basic relations. First, we introduce two notations:

$$\Sigma_{\leq r, \widehat{X}} \stackrel{\text{def}}{=} \{\mathbf{v} \in A \mid \text{rank}_X(\mathbf{v}) \leq r\},$$

$$\Sigma_{r, \widehat{X}} \stackrel{\text{def}}{=} \{\mathbf{v} \in A \mid \text{rank}_X(\mathbf{v}) = r\}.$$

Definition 2.2 (Maximal rank). The maximal X -rank is defined as the smallest r such that $\Sigma_{\leq r, \widehat{X}} = A$, and denoted by r_{\max} .

Definition 2.3. A rank r is called typical if $\Sigma_{r, \widehat{X}}$ contains an open Euclidean ball in A .

Remark 4. Since $\Sigma_{r, \widehat{X}}$ is a semialgebraic set [32], a rank r is typical if and only if $\Sigma_{r, \widehat{X}}$ has nonzero Lebesgue measure. Hence, a rank is typical, if and only if it appears with nonzero probability (if the vectors of A are drawn from an absolutely continuous probability distribution).

The following properties of typical ranks over \mathbb{C} and \mathbb{R} are known.

Lemma 2.4. If $\mathbb{K} = \mathbb{C}$, there exists only one typical rank, which is called generic rank, and denoted by r_{gen} . Moreover, the elements of rank r_{gen} are Zariski-dense in A , i.e. there exists an algebraic subvariety $Z \subsetneq A$ such that $\text{rank}_X(\mathbf{v}) = r_{\text{gen}}$ for any $\mathbf{v} \in A \setminus Z$.

Theorem 2.5 ([4]). Over the real field, the typical ranks form a contiguous set, i.e. there exist the numbers $r_{\text{typ}, \min}$ and $r_{\text{typ}, \max}$ such that:

- Any r_1 such that $r_{\text{typ}, \min} \leq r_1 \leq r_{\text{typ}, \max}$ is typical;
- Any r_1 such that $r_1 < r_{\text{typ}, \min}$ or $r_1 > r_{\text{typ}, \max}$ is not typical.

Next, the following theorem relates maximal and typical/generic ranks.

Theorem 2.6 ([6]).

- If $\mathbb{K} = \mathbb{R}$, then $r_{\max} \leq 2r_{\text{typ},\min}$.
- If $\mathbb{K} = \mathbb{C}$, then $r_{\max} \leq 2r_{\text{gen}}$.

Finally, there is a relation between real typical ranks and generic complex ranks.

Theorem 2.7 ([6]). Let $\hat{X}_{\mathbb{R}} = \hat{X}$ be a real variety satisfying Assumptions 1 to 5, and $\hat{X}_{\mathbb{C}} = \hat{X}_{\mathbb{R}} \otimes \mathbb{C}$ be its complexification. Then it holds that

$$r_{\text{typ},\min}(\hat{X}_{\mathbb{R}}) = r_{\text{gen}}(\hat{X}_{\mathbb{C}}),$$

i.e. the smallest typical real rank is equal to the complex generic rank.

Remark 5. All the varieties that we consider in this paper satisfy Assumptions 1 to 5.

2.4. Secant varieties and border rank. The r -th secant variety⁸ is, by definition, the Zariski closure of the elements of rank $\leq r$:

$$\sigma_r(\hat{X}) \stackrel{\text{def}}{=} \overline{\Sigma_{\leq r, \hat{X}}} \subseteq A.$$

The following properties of $\sigma_r(\hat{X})$ are known [27].

Lemma 2.8.

- If $\mathbb{K} = \mathbb{C}$, then $\sigma_r(\hat{X})$ is the Euclidean closure of $\Sigma_{\leq r, \hat{X}}$.
- If $\mathbb{K} = \mathbb{C}$, and $\dim \sigma_{r-1}(\hat{X}) < \dim \sigma_r(\hat{X})$, then a general point in $\hat{\sigma}_r(\hat{X})$ has rank r , i.e. there exist a subvariety $Y \subsetneq \hat{\sigma}_r(\hat{X})$, such that

$$\sigma_r(\hat{X}) \setminus Y \subset \Sigma_{r, \hat{X}}.$$

- If $\mathbb{K} = \mathbb{R}$, it is not the case: there may exist a nonempty open subset of $\sigma_r(\hat{X})$ such that each point in this open subset has X -rank strictly larger than r .

Nevertheless, there is a correspondence between real and complex varieties.

Lemma 2.9 ([32]). Let $\hat{X}_{\mathbb{R}} = \hat{X}$ be a real variety satisfying Assumptions 1 to 5, and $\hat{X}_{\mathbb{C}} = \hat{X}_{\mathbb{R}} \otimes \mathbb{C}$. Then for all r the secant variety $\sigma_r(\hat{X}_{\mathbb{R}})$ satisfies Assumptions 1 to 5, and $\sigma_r(\hat{X}_{\mathbb{C}})$ is a complexification of $\sigma_r(\hat{X}_{\mathbb{R}})$.

2.5. Defectivity, expected dimension and generic rank. In this subsection, we only consider the case $\mathbb{K} = \mathbb{C}$, and we assume that \hat{X} satisfies Assumptions 1 to 4.

Lemma 2.10. The dimensions of $\sigma_r(\hat{X})$ are increasing until $r = r_{\text{gen}}$, i.e.,

$$\begin{aligned} \dim \hat{X} = \dim \sigma_1(\hat{X}) &< \dim \sigma_2(\hat{X}) < \cdots < \dim \sigma_{r_{\text{gen}}-1}(\hat{X}) \\ &< \dim \sigma_{r_{\text{gen}}}(\hat{X}) = \dim \sigma_{r_{\text{gen}}+1}(\hat{X}) = \cdots = \dim A. \end{aligned}$$

theorem 2.10 is a direct consequence of Lemma 2.8 and tells us that we are able to find the generic rank by looking at dimensions of $\sigma_r(\hat{X})$. For this, a useful concept, i.e., the expected dimension, is introduced.

⁸Here we again prefer using affine varieties. For projective definitions, we invite the reader to consult [27].

Definition 2.11 (Expected dimension). The expected dimension of $\sigma_r(\widehat{X})$ is defined as

$$\exp \dim \sigma_r(\widehat{X}) \stackrel{\text{def}}{=} \min\{r \dim \widehat{X}, \dim A\}$$

The intuition behind theorem 2.11 is that if we add in (2.1) vectors from the variety of dimension $\dim \widehat{X}$, we obtain an object of dimension r times larger. In general, the following holds true.

Remark 6.

- In general, $\exp \dim \sigma_r(\widehat{X}) \geq \dim \sigma_r(\widehat{X})$.
- If there is a strict inequality, $\sigma_r(\widehat{X})$ is called defective. Otherwise $\sigma_r(\widehat{X})$ is called non-defective.

Corollary 2.12. The following bound on r_{gen} can be given:

$$r_{\text{gen}} \geq \left\lceil \frac{\dim A}{\dim \widehat{X}} \right\rceil \quad (2.2)$$

In particular, if all $\sigma_r(\widehat{X})$ are non-defective, then $r_{\text{gen}} = \left\lceil \frac{\dim A}{\dim \widehat{X}} \right\rceil$.

Example 6. The Alexander-Hirschowitz theorem [1] states that for $\widehat{X} = \nu_d(\mathbb{C}^m)$, all the secant varieties are non-defective except a finite number of exceptions. Hence, by Corollary 2.12 and table 2.1, the generic rank r_{gen} is equal to $\lceil r_1(m, d) \rceil$, where

$$r_1(m, d) \stackrel{\text{def}}{=} \frac{\binom{m+d-1}{d}}{m},$$

except $(m, d) \in \{(3, 3), (4, 3), (4, 5), (4, 6)\}$, where r_{gen} is increased by 1.

3. Uniqueness and identifiability.

3.1. Uniqueness of a decomposition. First, we introduce the notion of uniqueness.

Definition 3.1. An X -rank decomposition (2.1) is unique if all the other decompositions of the form (2.1) differ only by permutation of the summands in (2.1).

Remark 7. This definition corresponds to the standard definition of uniqueness of tensor decompositions. For instance, a tensor decomposition

$$\mathcal{T} = \mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1 + \mathbf{a}_2 \otimes \mathbf{b}_2 \otimes \mathbf{c}_2 \quad (3.1)$$

is unique if it is unique up to permutation of summands and exchange of scaling in the vectors.

In this paper, we study the notion of generic uniqueness, or uniqueness of “almost all” decompositions. The following algebraic definition is often adopted in the literature.

Definition 3.2. A variety $\widehat{X} \subset A$ is called r -identifiable if a general element in $\Sigma_{r, \widehat{X}}$ has a unique rank- r decomposition, i.e. there exists a semialgebraic subset $Z \subsetneq \Sigma_{r, \widehat{X}}$ of strictly smaller dimension such that any element in $\Sigma_{r, \widehat{X}} \setminus Z$ has a unique rank- r decomposition.

First, we remark on the relation between real and complex identifiability.

Lemma 3.3 ([32]). Assume that \widehat{X} satisfies Assumptions 1 to 5, $r < r_{\text{gen}}$ and $\widehat{X}_{\mathbb{C}}$ is r -identifiable. Then $\widehat{X}_{\mathbb{R}}$ is also r -identifiable.

In the remainder of the section, we consider only the complex case.

Remark 8. If $\mathbb{K} = \mathbb{C}$, \hat{X} satisfies Assumptions 1 to 4 and is r -identifiable, then any $\mathbf{v} \in \sigma_r(\hat{X})$ is a limit of a sequence of vectors $\mathbf{v}_k \in \Sigma_{r,\hat{X}}$ with a unique decomposition.

Next, we give some interpretation to theorem 3.2.

Lemma 3.4. Let $\mathbb{K} = \mathbb{C}$, \hat{X} satisfy Assumptions 1 to 4. Then \hat{X} is r -identifiable if and only if

$$\text{for } r \text{ general points } \mathbf{p}_1, \dots, \mathbf{p}_r \in \hat{X}, \mathbf{p}_1 + \dots + \mathbf{p}_r \text{ has a unique rank-}r \text{ decomposition.} \quad (3.2)$$

theorem 3.4 states that \hat{X} is r -identifiable if for “randomly chosen” $\mathbf{p}_1, \dots, \mathbf{p}_r \in \hat{X}$ their sum has a unique X -rank decomposition. The proof is given in section 6.1. This conclusion can be also transferred to the parameter space.

Corollary 3.5. Let \hat{X} be an algebraic variety over \mathbb{C} satisfying Assumptions 1 to 4, and such that there exists a polynomial map $\mathcal{X} : \mathbb{C}^M \rightarrow A$ such that $\hat{X} = \mathcal{X}(\mathbb{C}^M)$. Then \hat{X} is r -identifiable if and only if for a general point $(\mathbf{z}_1, \dots, \mathbf{z}_r) \in (\mathbb{C}^M)^{\times r}$, the decomposition

$$v = \mathcal{X}(\mathbf{z}_1) + \dots + \mathcal{X}(\mathbf{z}_r) \quad (3.3)$$

is unique, i.e., the semialgebraic set

$$Y = \{(\mathbf{z}_1, \dots, \mathbf{z}_r) \in (\mathbb{C}^M)^{\times r} \mid \mathbf{v} \text{ in eq. (3.3) has nonunique decompositions}\}.$$

has Lebesgue measure zero.

Proof. By [25, Exercise II 3.22], each \mathbf{z}_i is general in \mathbb{C}^M if and only if $\mathcal{X}(\mathbf{z}_i)$ is general in \hat{X} . Then the statement follows from Lemma 3.4. ■

Example 7. Consider the case of remark 7. The Segre variety $\text{Seg}(\mathbb{C}^{I_1} \times \mathbb{C}^{I_2} \times \mathbb{C}^{I_3})$ is 2-identifiable if and only if the decomposition eq. (3.1) is unique for general $\mathbf{a}_1, \mathbf{b}_1, \mathbf{c}_1, \mathbf{a}_2, \mathbf{b}_2, \mathbf{c}_2$ (i.e. drawn randomly with respect to an absolutely continuous probability distribution). Note the decomposition eq. (3.1) is unique does not mean $\mathbf{a}_1, \dots, \mathbf{b}_2$ are unique, in fact they are unique up to scaling.

Definition in the parameter space is more common in linear algebra and engineering literature. Hence theorem 3.5 establishes correspondence between these two definitions.

3.2. Necessary and sufficient conditions for generic uniqueness. Here, in what follows, we consider only the case $\mathbb{K} = \mathbb{C}$. First, the following result can be extended from [38].

Lemma 3.6 ([38]).

- If $\sigma_r(\hat{X})$ is defective, then \hat{X} is not r -identifiable.
- If $\sigma_r(\hat{X})$ is non-defective, then a general point in $\sigma_r(\hat{X})$ has a finite number of decompositions.

Thus, already looking at the dimension of $\sigma_r(\hat{X})$ we can already conclude that \hat{X} is r -identifiable. This can be done numerically using the Terracini’s lemma.

Lemma 3.7 (Terracini). Assume that \hat{X} satisfies Assumptions 1 to 4. Then for a general point $\mathbf{v} = \mathbf{p}_1, \dots, \mathbf{p}_r \in \sigma_r(\hat{X})$, the tangent space is

$$T_{\mathbf{v}}\sigma_r(\hat{X}) = \text{Span}\{T_{\mathbf{p}_1}\hat{X}, \dots, T_{\mathbf{p}_r}\hat{X}\}.$$

Hence, the non-defectivity can be checked numerically, by picking r “random” points and comparing $\dim T_{\mathbf{v}}\sigma_r(\widehat{X})$ with $\exp \dim \sigma_r(\widehat{X})$.

Definition 3.8. A variety \widehat{X} is called r -weakly defective if for r general points in \widehat{X} a general hyperplane tangent to them is tangent to X elsewhere [11]. If X is not r -weakly defective, then X is r -identifiable (the converse is not true).

3.3. Examples: Veronese and Segre-Veronese varieties. We review here some results on identifiability of varieties from table 2.1, that will be needed. First, recall a recent result that for all subgeneric ranks, the Veronese variety is r -identifiable.

Theorem 3.9 ([14, Theorem 1.1]). Let $d \geq 3$ and $m \geq 2$. Then $\nu_d(\mathbb{C}^m)$ is r -identifiable for all $r < r_2(m, d)$, where

$$r_2(m, d) = \begin{cases} r_1(m, d) - 1, & \text{if } (m, d) \in \{(4, 4), (3, 6), (6, 3)\}, \\ r_1(m, d), & \text{otherwise.} \end{cases} \quad (3.4)$$

Next, we recall stronger results on r -weak defectivity of the Veronese varieties.

Theorem 3.10 ([3, 30, 12]). Let $d \geq 3$ and $m \geq 2$. Then the Veronese variety $\nu_d(\mathbb{C}^m)$ is not r -weakly defective⁹ for $r < r_3(m, d)$, where

$$r_3(m, d) = \begin{cases} r_1(m, d) - \frac{m-2}{3}, & d = 3, \\ r_1(m, d), & \text{otherwise.} \end{cases}$$

For Segre-Veronese varieties, we are not aware of explicitly available results on identifiability. However, the identifiability of such varieties can be easily deduced from theorem 3.10 and the results of [7] on identifiability of Segre products of varieties.

Denote

$$r_4(m, n, d) = \frac{\binom{m+d-1}{d}}{m+n-1}. \quad (3.5)$$

Corollary 3.11. Let $m = \dim V \geq 2$, $d \geq 3$, $n = \dim W \geq 1$, and $k < r_4(m, n, d)$, where

$$r_5(m, n, d) = \begin{cases} r_2(m, d), & \text{if } n = 1, \\ \min(r_4(m, n, d), r_3(m, d)), & \text{if } n > 1. \end{cases} \quad (3.6)$$

Then the variety $\text{Seg}(\nu_d(V) \times W)$ is kn -identifiable.

Proof. The proof is given in section 6.1. ■

Remark 9. Although the expression in (3.6) looks complicated, in fact,

$$r_5(m, n, d) = r_4(m, n, d)$$

if $n > 1, d \geq 3$ or if $n = 1, (m, d) \notin \{(4, 4), (3, 6), (6, 3)\}$.

4. Veronese scrolls. In this section, we recall a variety that is a generalization of the well-known rational normal scroll [9].

⁹The case $2 \leq m \leq 3$ was proved in the proof of [12, Thm 5.1], $d = 3$ was proved in [30, Thm. 4.1], the case $d \geq 4$ is proved in [3, Thm. 1.1.] (see also [30, Corollary 4.5]).

4.1. Simultaneous Waring decompositions. Let $0 \leq a_1 \leq \dots \leq a_d$ be a sequence of natural numbers¹⁰ put in one vector $\mathbf{a} = (a_1, \dots, a_d) \in \mathbb{N}^d$ and define a shorthand notation

$$S^{\mathbf{a}}V \stackrel{\text{def}}{=} S^{a_1}V \oplus S^{a_2}V \oplus \dots \oplus S^{a_d}V,$$

which is a vector space of dimension

$$\dim(S^{\mathbf{a}}V) = \sum_{k=1}^d \binom{m + a_k - 1}{a_k}.$$

We say that $f = (f^{(1)}, \dots, f^{(d)}) \in S^{\mathbf{a}}V$ has a Waring-like decomposition of rank r if there exist $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $c_{k,l} \in \mathbb{K}$ such that

$$\begin{aligned} f^{(1)} &= c_{1,1}\mathbf{v}_1^{a_1} + \dots + c_{1,r}\mathbf{v}_r^{a_1}, \\ &\vdots \\ f^{(d)} &= c_{d,1}\mathbf{v}_1^{a_d} + \dots + c_{d,r}\mathbf{v}_r^{a_d}, \end{aligned} \tag{4.1}$$

In other words, decomposition (4.1) is equivalent to simultaneous Waring decompositions with the same vectors but different coefficients.

Example 8. Let us show that example 2 is a special case of the Waring-like decomposition (4.1). Since $f(\mathbf{0}) = \mathbf{0}$ in (1.2), we have that

$$f(\mathbf{u}) = f^{(1)}(\mathbf{u}) + \dots + f^{(d)}(\mathbf{u}),$$

where $f^{(d)}(\mathbf{u})$ is the d -th degree homogeneous part of $f(\mathbf{u})$. Hence, if the polynomial f admits a decomposition (1.2), then all the homogeneous parts $f^{(d)}$ can be decomposed as

$$\begin{aligned} f^{(1)} &= c_{1,1}(\mathbf{v}_1^\top \mathbf{u}) + \dots + c_{1,r}(\mathbf{v}_r^\top \mathbf{u}), \\ f^{(2)} &= c_{2,1}(\mathbf{v}_1^\top \mathbf{u})^2 + \dots + c_{2,r}(\mathbf{v}_r^\top \mathbf{u})^2, \\ &\vdots \\ f^{(d)} &= c_{d,1}(\mathbf{v}_1^\top \mathbf{u})^d + \dots + c_{d,r}(\mathbf{v}_r^\top \mathbf{u})^d. \end{aligned}$$

which is a special case of eq. (4.1) for the vector of integers $\mathbf{a} = (1, \dots, d)$.

4.2. Veronese scrolls: a parametric definition. The decomposition eq. (4.1) can be put in the framework of X -rank as follows. Define the following map:

$$\begin{aligned} \psi : V \times \mathbb{K}^d &\rightarrow S^{\mathbf{a}}V \\ (\mathbf{v}, (c_1, \dots, c_d)) &\mapsto (c_1\mathbf{v}^{a_1}, c_2\mathbf{v}^{a_2}, \dots, c_d\mathbf{v}^{a_d}), \end{aligned} \tag{4.2}$$

and define the image of this map as

$$\widehat{X}_{\mathbf{a}} = \widehat{X}_{\mathbf{a},V} \stackrel{\text{def}}{=} \psi(V \times \mathbb{K}^d), \tag{4.3}$$

¹⁰By convention, \mathbb{N} is the set of nonnegative integers and includes 0.

and $X_{\mathbf{a}} = X_{\mathbf{a},V} \subset \mathbb{P}S^{\mathbf{a}}V$ the corresponding subset in the projective space.

Remark 10. *It is easy to see that $f = (f^{(1)}, \dots, f^{(d)}) \in S^{\mathbf{a}}V$ has a Waring-like decomposition if and only if it has an X -rank decomposition with $\hat{X} = \hat{X}_{\mathbf{a},V}$.*

It can be shown that $\hat{X}_{\mathbf{a},V}$ satisfies Assumptions 1 to 4 (affine cone of a projective variety $\hat{X}_{\mathbf{a},V}$). In particular,

- For $\dim(V) = m = 2$, $X_{\mathbf{a},V}$ is the rational normal (d -fold) scroll, a classic object in algebraic geometry [9].
- For $m \geq 2$, $X_{\mathbf{a},V}$ can be realized as a projective bundle¹¹ $X_{\mathbf{a},V} \simeq \mathbb{P}(\mathcal{O}_{\mathbb{P}V}(a_1) \oplus \dots \oplus \mathcal{O}_{\mathbb{P}V}(a_d))$ [2, 9, 17].

Instead, in one of the following sections, we give explicit (ideal-theoretic) defining equations for the set eq. (4.3), which will provide an alternative proof that $\hat{X}_{\mathbf{a},V}$ is a variety.

Remark 11. *For $m > 2$, the object $X_{\mathbf{a},V}$ does not have a name in the literature, and we call it Veronese scroll, as a hybrid of “rational normal scroll” and “Veronese variety”.*

Remark 12. *Now consider the following map.*

$$\begin{aligned} \psi_m : W \times V \times \mathbb{K}^d &\rightarrow S^{\mathbf{a}}V \otimes W \\ &(\mathbf{w}, \mathbf{v}, (c_1, \dots, c_d)) \mapsto \begin{pmatrix} w_1 c_1 \mathbf{v}^{a_1}, & b_1 c_2 \mathbf{v}^{a_1}, & \dots & w_1 c_d \mathbf{v}^{a_d}, \\ \vdots & & & \vdots \\ w_n c_1 \mathbf{v}^{a_1}, & b_m c_2 \mathbf{v}^{a_2}, & \dots & w_n c_d \mathbf{v}^{a_d} \end{pmatrix}, \end{aligned}$$

and define $\hat{Y}_{(a_1, \dots, a_d)}$ the image of ψ_m . It is easy to see that $\hat{Y}_{(a_1, \dots, a_d)} = \text{Seg}(\hat{X}_{(a_1, \dots, a_d)} \times W)$. Moreover, as in section 4.1, we can show that the polynomial decomposition eq. (1.1) is exactly the X -rank decomposition for $\hat{Y}_{(1, \dots, d)}$.

4.3. Determinantal construction (defining equations). This section is not needed to prove the main results of the paper, but still gives more insight in the nature of the Veronese scrolls.

First, recall a definition of the catalecticant matrix [26, Ch. 1] (we prefer giving it in coordinates). Let $f \in S^d V$ be given by coordinates $\{f_{\alpha}\}_{\alpha \in \Delta_{s,m}}$, as defined in appendix A.1.

Then the first catalecticant matrix, for $1 \leq s \leq d$, is defined as¹²

$$C_f \in \mathbb{K}^{m \times \binom{m+d-2}{d-1}}, \text{ where } (C_f)_{i,\beta} = f_{(\beta_1, \dots, \beta_i+1, \dots, \beta_m)},$$

where the columns are indexed by $\beta \in \Delta_{s,m-1}$.

Proposition 4.1. *Let $a_k \geq 1$, and $f = (f^{(1)}, \dots, f^{(d)}) \in S^{\mathbf{a}}V$. Define the stacked matrix as*

$$S(f) \stackrel{\text{def}}{=} [C_{f_1} \mid \dots \mid C_{f_d}]. \quad (4.4)$$

Then it holds that

$$f \in \hat{X}_{\mathbf{a},V} \iff \text{rank } S(f) \leq 1,$$

i.e. $\hat{X}_{\mathbf{a},V}$ is defined (set-theoretically) by the vanishing of all 2×2 minors of $S(f)$.

Proof. The proof is contained in section 6.2 ■

¹¹We are not reproducing the bundle construction, since it is difficult without going into technical details.

¹²In fact, this is the matrix representation map $S^{d-s}V^* \rightarrow S^s V^*$ given by differentiation.

Remark 13. A similar construction for the matrix $S(f)$ can be found in [2, §3].

Proposition 4.2. Let $a_k \geq 1$, and $S(f)$ be defined as in eq. (4.4). Then the 2×2 minors of $S(f)$ generate the ideal of $\widehat{X}_{\mathbf{a},V}$. The proposition is much stronger than theorem 4.1. The proof relies on the tools of representation theory, and is contained in section 6.2.

5. Main results. Throughout this section we assume that $\mathbb{K} = \mathbb{C}$. By [32, Section 5], all our results hold for the real case too. We will also use a shorthand $X_{\mathbf{a}}$ instead of $X_{\mathbf{a},V}$.

5.1. Identifiability of Veronese scrolls and polynomial decompositions. **Proposition 5.1.** Let $m = \dim V \geq 2$, $a_d \geq 3$, $n = \dim W \geq 1$. Next, consider the Veronese scroll $\widehat{X}_{(a_1, \dots, a_d)}$ with $\mathbf{a} = (a_1, \dots, a_d)$, $1 \leq a_1 \leq \dots \leq a_d$, and the variety $\widehat{Y} = \widehat{Y}_{(a_1, \dots, a_d)}$. Then we have the following.

1. \widehat{Y} is r -identifiable if

$$r \leq \min(\lceil r_5(m, n, a_d) \rceil - 1, \dim S^{a_1} V)n. \quad (5.1)$$

2. \widehat{Y} cannot be r -identifiable for $r > n \dim(S^{a_1} V)$.

The proof is given in section 6.3, and the idea of the proof is based on two facts:

1. Under the condition (5.1), the highest degree terms are generically unique, and \mathbf{w}_k and \mathbf{v}_k are uniquely determined.
 2. The lower degree terms (coefficients $c_{k,l}$) can be recovered using a simple linear algebra.
- theorem 5.1 has immediate implications for the polynomial decomposition (1.1), which corresponds to the case where degrees are defined by $\mathbf{a} = (1, \dots, d)$.

Corollary 5.2. Let d, m, n be such that $d \geq 3$, $m \geq 2$, and consider the field \mathbb{C}

1. The decomposition (1.1) is r -identifiable if

$$r \leq \min(m, \lceil r_5(m, n, d) \rceil - 1) \cdot n. \quad (5.2)$$

In particular, if $m < r_5(m, n, d)$, then the model (1.1) is mn -identifiable.

2. The model (1.1) cannot be r -identifiable for $r > mn$.

First, let us give some examples. In Tables 5.1a and 5.2a, we provide the calculated bound on identifiability (5.2) for $d = 3, 4$. For comparison, we show the maximal non-defective rank obtained numerically¹³ using theorem 3.7. In all tables, the cases when the rank coincides with mn (i.e., the maximal possible rank by theorem 5.2, part 2) are shown in bold.

As it is easy to see from tables 5.1a and 5.2b, that the bound given by (5.2) does not detect the maximum identifiability bound obtained by Terracini's lemma (especially for $m < n$), but does perform well for the case $m \geq n$. Moreover, the following remark can be made.

Corollary 5.3.

1. For fixed m and n , there exists d_0 such that the inequality $m < r_5(m, n, d)$ holds for all $d \geq d_0$.
2. If $m \geq n \geq 4$, then $m < r_5(m, n, d)$ holds true for all $d \geq 4$.

Proof.

1. This fact follows since the numerator of (3.5) for $m > 1$ is a strictly increasing in d .

¹³We also checked that the weak tangential nondefectivity described in [13] holds for all cases Tables 5.1b and 5.2b, except when $r = mn$ (in that case, the weak tangential nondefectivity criterion works up to $mn - 1$).

$\begin{smallmatrix} n \\ m \end{smallmatrix}$	1	2	3	4	5	6	7	8
2	1	2	0	0	0	0	0	0
3	3	4	3	4	5	6	7	0
4	4	6	9	8	10	12	7	8
5	5	10	12	16	15	18	21	16
6	6	12	18	24	25	30	28	32
7	7	14	21	28	35	36	42	40
8	8	16	24	32	40	48	56	56

(a) Our bound

$\begin{smallmatrix} n \\ m \end{smallmatrix}$	1	2	3	4	5	6	7	8
2	2	3	3	4	5	6	6	6
3	3	6	7	9	10	11	12	12
4	4	8	12	15	17	18	19	20
5	5	10	15	20	25	27	29	31
6	6	12	18	24	30	36	41	44
7	7	14	21	28	35	42	49	56
8	8	16	24	32	40	48	56	64

(b) Terraccini lemma

Table 5.1
Case $d = 3$.

$\begin{smallmatrix} n \\ m \end{smallmatrix}$	1	2	3	4	5	6	7	8
2	2	2	3	0	0	0	0	0
3	3	6	6	8	10	6	7	8
4	4	8	12	16	20	18	21	24
5	5	10	15	20	25	30	35	40
6	6	12	18	24	30	36	42	48
7	7	14	21	28	35	42	49	56
8	8	16	24	32	40	48	56	64

(a) Our bound

$\begin{smallmatrix} n \\ m \end{smallmatrix}$	1	2	3	4	5	6	7	8
2	2	4	6	6	6	6	7	8
3	3	6	9	12	15	18	19	20
4	4	8	12	16	20	24	28	32
5	5	10	15	20	25	30	35	40
6	6	12	18	24	30	36	42	48
7	7	14	21	28	35	42	49	56
8	8	16	24	32	40	48	56	64

(b) Terraccini's lemma

Table 5.2
Case $d = 4$.

2. If $5 \geq m \geq n \geq 4$, this can be verified from Table 5.2a. If $d \geq 4$ and $m \geq n > 1$, $m \geq 5$, then

$$\frac{r_5(m, n, d)}{m} \geq \frac{r_5(m, n, 4)}{m} = \frac{(m+1)(m+2)(m+3)}{2 \cdot 3 \cdot 4 \cdot (m+n-1)} > 1.$$

■

Remark 14. The authors in [20] suggest the bound

$$(m-1)m(n-1)n \geq 2(r-1)r \quad (5.3)$$

for decomposition (1.1), also shown in table 5.3a. The bound eq. (5.3) appears from Kruskal-type generic uniqueness conditions for unstructured $m \times n \times N$ tensors [19]. In fact, a better bound exists for unbalanced tensors, which is $(m-1)(n-1)$ [13].

We make two remarks here:

1. The bound mn is better than the heuristic bound (5.3) (see the values Table 5.3a).
2. The tensor considered in [20] is structured, and the bound (5.3) cannot be directly

applied to model (1.1)¹⁴. In fact, for degree 2 (see table 5.3b), the model can be non-identifiable even if the bound (5.3) holds.

$\begin{smallmatrix} n \\ m \end{smallmatrix}$	1	2	3	4	5	6	7	8
2	1	1	2	3	3	4	5	5
3	1	2	3	4	6	7	8	9
4	1	3	4	6	8	10	11	13
5	1	3	6	8	10	12	15	17
6	1	4	7	10	12	15	18	21
7	1	5	8	11	15	18	21	24
8	1	5	9	13	17	21	24	28

$\begin{smallmatrix} n \\ m \end{smallmatrix}$	1	2	3	4	5	6	7	8
2	1	2	3	3	3	3	3	3
3	1	3	4	4	5	6	6	6
4	1	4	6	7	7	8	8	9
5	1	5	7	8	10	10	11	12
6	1	6	9	10	12	13	14	15
7	1	7	10	12	14	16	17	18
8	1	8	12	14	16	18	20	22

(a) The heuristic bound given in (5.3).

(b) Our bound

Table 5.3Case $d = 2$.

In fact even if the model is non-identifiable, the decomposition can be partially unique.

Corollary 5.4. Let s be a number $1 < s < d$ such that $\binom{m+s-1}{s} < r_5(m, n, d)$. Then for all $r \leq \binom{m+s-1}{s}n$, the decomposition (1.1) is partially identifiable except the terms of degree less than s . That is, all the elements in the decomposition (1.1) can be determined uniquely (up to trivial indeterminacies), except the coefficients $c_{k,l}$, for $k < s$.

5.2. Dimensions of secant varieties. From Proposition 5.2 we can immediately find dimensions of secant varieties for small ranks.

Proposition 5.5. Let $m = \dim V \geq 2$, $n = \dim W \geq 1$, and $\mathbf{a} = (a_1, \dots, a_d)$, $1 \leq a_1 \leq \dots \leq a_{d-1} < a_d$, with $a_d \geq 3$. Consider the variety $\hat{Y} = \hat{Y}_{(a_1, \dots, a_d)}$, and assume that

$$r \leq (\lceil r_5(m, n, a_d) \rceil - 1)n.$$

Then we have that:

1. If $r \leq n \dim S^{a_1} V$, then \hat{Y} is non-defective, i.e.

$$\dim \sigma_r(\hat{Y}) = \exp \dim \sigma_r(\hat{Y}) = r(m + n + d - 2),$$

2. If $n \dim S^{a_s} V < r \leq n \dim S^{a_{s+1}} V$ then

$$\sigma_r(\hat{Y}) = \left(S^{(a_1, \dots, a_s)} V \otimes W \right) \times \sigma_r(\hat{Y}_{(a_{s+1}, \dots, a_d)}), \quad (5.4)$$

and hence

$$\dim \sigma_r(\hat{Y}) = \exp \dim \sigma_r(\hat{Y}_{(a_{s+1}, \dots, a_d)}) + \dim \left(S^{(a_1, \dots, a_s)} V \otimes W \right), \quad (5.5)$$

$$= r(m + n + d - 2 - s) + n \left(\sum_{j=1}^s \dim S^{a_j} V \right). \quad (5.6)$$

¹⁴Take for instance the simple case of symmetry. The maximal symmetric rank R_s^o for which symmetric tensors will have a unique CP decomposition is smaller [14] than the maximal rank R^o for which unconstrained tensors will have a unique CP decomposition [15, 13].

The proof is based on theorem 5.2, and is contained in section 6.3.

It may be easier to look at the dimensions in terms of so-called defects of \hat{Y} , defined as

$$\delta_r(Y) \stackrel{\text{def}}{=} \exp \dim \sigma_r(\hat{Y}) - \dim \sigma_r(\hat{Y}),$$

where $\delta_r(\hat{Y})$ is called the defect of $\sigma_r(\hat{Y})$. Then Proposition 5.5 can be reformulated as follows.

Proposition 5.6 (Proposition 5.5 reformulated.) *Under the assumptions of Proposition 5.5, the defect can be expressed as*

$$\delta_r(\hat{Y}) = \sum_{j=1}^d \max(r - n \dim S^{a_j} V, 0).$$

5.3. Generic ranks. In this section, we consider only the case $n = 1$, and $\mathbf{a} = (1, \dots, d)$. From theorem 5.5 it follows that the behaviour of the ranks of secant varieties depends only on higher degrees. As shown by the next lemma, for fixed d and large m everything depends on two higher degrees.

Lemma 5.7. *Let $d \geq 3$.*

1. *For all $m \geq 2$, it holds that $r_1(m, d) < \dim S^{d-1} V$.*
2. *For all $m > (d-2)(d-1)$ it holds that $\dim S^{d-2} V < r_1(m, d)$.*

Proof.

1. First, for $d \geq 2$ and $m \geq 2$ it holds that $m + d - 1 < md$. Therefore,

$$r_1(m, d) = \binom{m+d-2}{d-1} \frac{m+d-1}{md} < \binom{m+d-2}{d-1}.$$

2. As in the previous item, we have that

$$\frac{r_1(m, d)}{\binom{m+d-3}{d-2}} = \frac{(m+d-2)(m+d-1)}{m(d-1)d} > 1$$

The ratio is greater than one since $m \geq 2$ and $m > (d-2)(d-1)$.

From theorem 5.5 and theorem 5.7, we have the following immediate corollary.

Corollary 5.8. *Under the conditions 1–2 in theorem 5.7, we have*

$$r_{\text{gen}}(\hat{X}_{(1, \dots, d)}) = r_{\text{gen}}(\hat{X}_{(d-1, d)}).$$

The main result in this subsection is on the bound on generic rank of $\hat{X}_{(d-1, d)}$.

Proposition 5.9. *Let $d \geq 4$ and $m > 5$. Then*

$$\left\lceil \frac{\binom{m+d-2}{d-1} + \binom{m+d-1}{d}}{m+1} \right\rceil \leq r_{\text{gen}}(\hat{X}_{(d-1, d)}) \leq \left\lceil \frac{\binom{m+d-2}{d-1} + (m-1) \lceil \frac{\binom{m+d-1}{d}}{m} \rceil}{m} \right\rceil$$

The lower bound just follows from theorem 2.12, the whole proof is given in section 6.4. For large m , the lower bound is exact.

Proposition 5.10. *Let $d \geq 4$ and $m > (d-1)^2$. then*

$$r_{gen} = \left\lceil \frac{\binom{m+d-2}{d-1} + \binom{m+d-1}{d}}{m+1} \right\rceil.$$

In fact from the proof of Proposition 5.10, we can also obtain

Proposition 5.11. *When $d \geq 4$, $m > (d-1)^2$, and $r < r_{gen}$, $\widehat{X}_{(d-1,d)}$ is r -identifiable.*

As a corollary of theorem 5.10 and theorem 2.6 we obtain the following bound on r_{max} for polynomial decomposition eq. (1.2).

Corollary 5.12. *Let $\mathbb{K} = \mathbb{C}$ or \mathbb{R} , and fix m and d such that $d \geq 4$ and $m > (d-1)^2$. Then the maximal rank for the decomposition eq. (1.2) is bounded by*

$$r_{max} \leq 2 \cdot \left\lceil \frac{\binom{m+d-2}{d-1} + \binom{m+d-1}{d}}{m+1} \right\rceil. \quad (5.7)$$

The bound in Corollary 5.12 implies that

$$r_{max} \leq 2 \left\lceil \binom{m+d-2}{d-1} \frac{m+2d-1}{(m+1)d} \right\rceil \leq \frac{2}{d} \binom{m+d-2}{d-1} \left(1 + \frac{2(d-1)}{(m+1)} \right). \quad (5.8)$$

Hence, the bound (5.8) is better than (1.6) if $m > 8$, and the ratio between the bounds (5.8) and (1.6) approaches $\frac{2}{d}$ asymptotically as $m \rightarrow \infty$.

6. Proofs.

6.1. Basic results on generic uniqueness. *Proof.* [Proof of theorem 3.4] Let

$$\text{Sec}_r^\circ(\widehat{X}) \stackrel{\text{def}}{=} \{(\mathbf{p}_1, \dots, \mathbf{p}_r, \mathbf{v}) \in \widehat{X}^{\times r} \times V : \mathbf{p}_1, \dots, \mathbf{p}_r \in \widehat{X}, \mathbf{v} = \mathbf{p}_1 + \dots + \mathbf{p}_r\},$$

and let $\pi_1: \widehat{X}^{\times r} \times V \rightarrow \widehat{X}^{\times r}$ and $\pi_2: \widehat{X}^{\times r} \times V \rightarrow V$ be the projections. Observe that $\sigma_r(\widehat{X})$ is the Zariski closure of $\pi_2(\text{Sec}_r^\circ(\widehat{X}))$, and $\pi_1: \text{Sec}_r^\circ(\widehat{X}) \rightarrow \widehat{X}^{\times r}$ is an isomorphism.

Then that $\sigma_r(\widehat{X})$ is identifiable implies $\pi_2: \text{Sec}_r^\circ(\widehat{X}) \rightarrow \sigma_r(\widehat{X})$ is birational, and thus the model \widehat{X} is r -identifiable in the sense of eq. (3.2). On the other hand, if the model \widehat{X} is r -identifiable in the sense of eq. (3.2), the cardinality of $\pi_2^{-1}\pi_2(p)$ is 1 for a general $p \in \text{Sec}_r^\circ(\widehat{X})$. Since $\pi_2(\text{Sec}_r^\circ(\widehat{X}))$ contains a Zariski dense open subset of $\sigma_r(\widehat{X})$, then $\pi_2: \text{Sec}_r^\circ(\widehat{X}) \rightarrow \sigma_r(\widehat{X})$ is birational, which implies $\sigma_r(\widehat{X})$ is identifiable. ■

In order to get results on identifiability of some Segre-Veronese varieties, we use a lemma that is a weaker version of the general result from [7].

Lemma 6.1 (Corollary of [7, Lemma 3.1, Corollary 3.3].) *Let $X \subset \mathbb{P}A$ is a smooth non-degenerate projective variety, and W be a vector space. Let $Y = \text{Seg}(X \times \mathbb{P}W)$ be the Segre embedding (such that $\dim(Y) = \dim(X) + \dim(W) - 1$). If X is not r -weakly defective, and*

$$r(\dim(Y) + 1) < \dim(A), \quad (6.1)$$

then Y is $(r \cdot \dim(W))$ -identifiable.

Proof. [Proof of theorem 3.11] For the case $n = 1$, this is just Theorem 3.9. Now we consider $n > 1$, and check the conditions of Lemma 6.1. In this case, the condition (6.1) is equivalent to $r < r_4(m, n, d)$. Since $r_5(m, n, d) = \min(r_4(m, n, d), r_3(m, d))$, the proof is complete. ■

6.2. Defining equations of Veronese scrolls. **Proof.** [Proof of theorem 4.1] \Rightarrow This direction is evident. In this case $(f^{(1)}, \dots, f^{(d)}) = (c_1 \mathbf{v}^{a_1}, \dots, c_d \mathbf{v}^{a_d})$. Since each f_k is rank-one, by [26, Thm. 1.28] each catalecticant matrix $C_{f^{(k)}}$ has rank ≤ 1 . Moreover the column space of each rank-one $C_{f^{(1)}}$ is spanned by the vector \mathbf{v} , therefore the column space of $S(f)$ is spanned by \mathbf{v} , and its rank does not exceed 1.

\Leftarrow Now consider $S(f)$ with rank 1 (the case of rank 0 is obvious). Define as \mathbf{v} the vector that spans the column space of $S(f)$. Since each of the matrices $C_{f^{(k)}}$ has rank ≤ 1 , from [26, Thm. 1.28] we have that $f = (f_1, \dots, f_d) = (c_1 \mathbf{v}_1^{a_1}, \dots, c_d \mathbf{v}_d^{a_d})$. But, from the apolarity [26, Ch. 1], [21], all the vectors \mathbf{v}_k must be collinear to \mathbf{v} . Therefore $f \in \hat{X}_{\mathbf{a}, V}$. ■

Since $\hat{X}_{\mathbf{a}, V}$ is invariant under the general linear group $\text{GL}(V)$, each degree- k component of the ideal $I(\sigma_r(\hat{X}_{\mathbf{a}, V}))$, denoted by $I_k(\sigma_r(\hat{X}_{\mathbf{a}, V}))$, in $S^k(S^{a_1}V \oplus \dots \oplus S^{a_d}V)$ is a representation of $\text{GL}(V)$. For any V ,

$$S^k(S^{a_1}V \oplus \dots \oplus S^{a_d}V) = \bigoplus_{l_1 + \dots + l_d = k} S^{l_1}(S^{a_1}V) \otimes \dots \otimes S^{l_d}(S^{a_d}V),$$

which is isomorphic to a direct sum of some irreducible representations $S_\mu V$ of $\text{GL}(V)$, where μ is a partition of $l_1 a_1 + \dots + l_d a_d$. Therefore, $I_k(\sigma_r(\hat{X}_{\mathbf{a}, V}))$ is isomorphic to a direct sum of some $S_\mu V$'s. Let $S_{\bar{\mu}}V$ denote a special realization of $S_\mu V$ in $S^k(S^{a_1}V \oplus \dots \oplus S^{a_d}V)$, see for example [27, Section 6] for more details. Similar to [28, Proposition 4.4] we have

Proposition 6.2. *Given vector spaces V, W with $r \leq \dim V \leq \dim W$ and $\dim V \geq 2$, then $S_{\bar{\pi}}V \subset I_k(\sigma_r(\hat{X}_{\mathbf{a}, V}))$ if and only if $S_{\bar{\pi}}W \subset I_k(\sigma_r(\hat{X}_{\mathbf{a}, W}))$.*

Proof. Given a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_{\dim V}\}$ for V , and a basis $\{\mathbf{w}_1, \dots, \mathbf{w}_{\dim W}\}$ for W , fix an embedding $i: V \hookrightarrow W$ such that $i(\mathbf{v}_j) = \mathbf{w}_j$ for $1 \leq j \leq \dim V$. Since each irreducible representation is generated by its highest weight vector, then

$$S_{\bar{\pi}}W = \text{GL}(W) \cdot S_{\bar{\pi}}V$$

for any π with length $\ell(\pi) \leq \dim V$ (See [23, 27]). The map i induces an embedding

$$\sigma_r(\hat{X}_{\mathbf{a}, V}) \xrightarrow{i} \sigma_r(\hat{X}_{\mathbf{a}, W}).$$

So in $S^k(S^{a_1}W \oplus \dots \oplus S^{a_d}W)$, we have $I_k(\sigma_r(\hat{X}_{\mathbf{a}, W})) \subset I_k(i(\sigma_r(\hat{X}_{\mathbf{a}, V})))$, which implies if $S_{\bar{\pi}}W \subset I_k(\sigma_r(\hat{X}_{\mathbf{a}, W}))$ then $S_{\bar{\pi}}V \subset I_k(\sigma_r(\hat{X}_{\mathbf{a}, V}))$.

Now we need to show for any $S_{\bar{\mu}}V \subset I_k(\sigma_r(\hat{X}_{\mathbf{a}, V}))$, $S_{\bar{\mu}}W \subset I_k(\sigma_r(\hat{X}_{\mathbf{a}, W}))$. Let

$$\sigma_r^\circ(\hat{X}_{\mathbf{a}, V}) := \{p \in \sigma_r(\hat{X}_{\mathbf{a}, V}) : p = (c_{1,1} \mathbf{u}_1^{a_1}, \dots, c_{d,1} \mathbf{u}_1^{a_d}) + \dots + (c_{1,r} \mathbf{u}_r^{a_1}, \dots, c_{d,r} \mathbf{u}_r^{a_d}), \quad (6.2)$$

$$\text{where } \mathbf{u}_1, \dots, \mathbf{u}_r \text{ are linearly independent}\}, \quad (6.3)$$

which is a Zariski dense open subset of $\sigma_r(\hat{X}_{\mathbf{a}, V})$. Since $I(\sigma_r(\hat{X}_{\mathbf{a}, V})) = I(\sigma_r^\circ(\hat{X}_{\mathbf{a}, V}))$, we only need to show for any $f \in S_{\bar{\mu}}W \subset I_k(i(\sigma_r^\circ(\hat{X}_{\mathbf{a}, V})))$, $f \in I_k(\sigma_r^\circ(\hat{X}_{\mathbf{a}, W}))$. But this is true due to

the fact $\mathrm{GL}(W) \cdot \sigma_r^\circ(\widehat{X}_{\mathbf{a},V}) = \sigma_r^\circ(\widehat{X}_{\mathbf{a},W})$. More precisely, for any $p \in \sigma_r^\circ(\widehat{X}_{\mathbf{a},W})$, since there is some $g \in \mathrm{GL}(W)$ such that $g \cdot p \in i(\sigma_r^\circ(\widehat{X}_{\mathbf{a},V}))$,

$$f(p) = f(g^{-1} \cdot g \cdot p) = (g^{-1} \cdot f)(g \cdot p) = 0,$$

which implies $f \in I_k(\sigma_r^\circ(\widehat{X}_{\mathbf{a},W}))$. ■

As a corollary of Proposition 6.2 we have

Proposition 6.3. *Given a vector space V with $2 \leq \dim V$, then*

$$S_{\overline{\pi}}V \subset I_k(\widehat{X}_{\mathbf{a},V}) \iff S_{\overline{\pi}}V \subset I_k(\widehat{X}_{\mathbf{a},\mathbb{C}^2}).$$

Since the ideal of the $\widehat{X}_{\mathbf{a},\mathbb{C}^2}$ is generated by 2×2 minors of $S(f)$ [22, Proposition 4.5], we conclude that theorem 4.2 is proved.

6.3. Identifiability and dimensions of secant varieties of Veronese scrolls. *Proof.* [Proof of theorem 5.1]

1. We have that

$$S^{\mathbf{a}}V \otimes W \simeq (S^{a_1}V \otimes W) \oplus \cdots \oplus (S^{a_d}V \otimes W),$$

and consider the j -th canonical projection $\pi_j : S^{\mathbf{a}}V \otimes W \rightarrow S^{a_j}V \otimes W$. Let $r = kn$.

Consider $\widehat{Y}_{(a_d)} = \pi_d(\widehat{Y})$. Then, by properties of Zariski closures, we have that $\sigma_r(\widehat{Y}_{(a_d)}) = \overline{\pi_d(\sigma_r(\widehat{Y}))}$, and by theorem A.9, a general point in $\sigma_r(\widehat{Y}_{(a_d)})$ belongs to $\pi_d(\widehat{\sigma}_r(Y))$. ■

Hence, we can take a general element

$$\mathbf{p} = (f_1^{(1)}, \dots, f_n^{(1)}, \dots, f_1^{(d)}, \dots, f_n^{(d)}) \in \widehat{\sigma}_r(Y)$$

such that $\mathbf{p} = \mathbf{y}_1 + \cdots + \mathbf{y}_r$, $\mathbf{y}_k \in \widehat{Y}$ and the decomposition

$$\pi_d(\mathbf{p}) = \pi_d(\mathbf{y}_1) + \cdots + \pi_d(\mathbf{y}_r),$$

is unique as X -rank decomposition with respect to $\widehat{Y}_{(a_d)}$ (due to r -identifiability of Z , which follows from theorem 3.11). A general $\mathbf{y}_l \in \widehat{Y}$, has the form

$$\mathbf{y}_l = (c_{1,l}\mathbf{v}_l^{a_1} \otimes \mathbf{w}_l, \dots, c_{d-1,l}\mathbf{v}_l^{a_{d-1}} \otimes \mathbf{w}_l, \mathbf{v}_l^{a_d} \otimes \mathbf{w}_l).$$

where the vectors $(\mathbf{v}_l \otimes \mathbf{w}_l)$ are determined uniquely, and $\{\mathbf{v}_l \otimes \mathbf{w}_l\}_{l=1}^r$ are linearly independent since $r \leq \dim(S^{a_1}V \otimes W)$.

Finally, the coefficients $c_{k,l}$ for $k < d$ should satisfy the equation

$$\pi_k(\mathbf{p}) = c_{k,1}\mathbf{v}_1^{a_k} \otimes \mathbf{w}_1 + \cdots + c_{k,r}\mathbf{v}_r^{a_k} \otimes \mathbf{w}_r. \quad (6.4)$$

By properties of Veronese embeddings, the vectors in $\{\mathbf{v}_l^{a_k} \otimes \mathbf{w}_l\}_{l=1}^r$ are also linearly independent, and therefore $c_{k,l}$ are determined uniquely.

2. Again, look at (6.4) for $k = 1$. We have that any system $\{\mathbf{v}_l^{a_1} \otimes \mathbf{w}_l\}_{l=1}^r$ is linearly dependent due to the fact that $r > \dim(S^{a_1}V \otimes W)$. Therefore, $Y_{(a_1, \dots, a_d)}$ cannot be r -identifiable.

Proof. [Proof of theorem 5.5]

1. By Proposition 5.1, $\sigma_r(\widehat{Y})$ is identifiable and thus nondefective.
2. It is sufficient to prove eq. (5.4), the rest follows automatically. Let $\pi = \pi_{(s,\dots,d)} : S^{(a_1,\dots,a_d)}V \otimes W \rightarrow S^{(a_{s+1},\dots,a_d)}V \otimes W$ denote the canonical projection *i.e.*,

$$\pi_{(s,\dots,d)} : (f_1^{(1)}, \dots, f_1^{(d)}, \dots, f_n^{(1)}, \dots, f_n^{(d)}) \mapsto (f_1^{(s)}, \dots, f_1^{(d)}, \dots, f_n^{(s)}, \dots, f_n^{(d)}).$$

As in the proof of Proposition 5.1, we have that $\sigma_r(\widehat{Z}) = \overline{\pi(\sigma_r(\widehat{Y}))}$ and by theorem A.9 a general element $\sigma_r(\widehat{Z})$ can be taken from $\pi(\sigma_r(\widehat{Y}))$.

Next, as in Proposition 5.1, there exists a Zariski-open subset of $U \subset \pi(\sigma_r(\widehat{Y}))$ such that any $\widehat{u} \in U$ has the decomposition $\widehat{u} = \widehat{u}_1 + \dots + \widehat{u}_r$, where

$$\widehat{u}_l = (c_{s+1,l} \mathbf{v}_l^{a_{s+1}} \otimes \mathbf{w}_l, \dots, c_{d-1,l} \mathbf{v}_l^{a_{d-1}} \otimes \mathbf{w}_k, \mathbf{v}_l^{a_d} \otimes \mathbf{w}_l),$$

and $v_l \otimes w_l$ are in general position. Therefore, for any $\widehat{p} \in S^{(a_1,\dots,a_d)}V$ with $\pi(\widehat{p}) = \widehat{u}$ and all $k \leq s$, the equation (6.4) is always solvable. Thus we have that

$$\pi^{-1}(U) = (S^{(a_1,\dots,a_s)}V \otimes W) \times U \subset \sigma_r(\widehat{Y}),$$

and, moreover, $\sigma_r(\widehat{Z}) = \overline{\pi^{-1}(U)} = \overline{(S^{(a_1,\dots,a_s)}V \otimes W) \times U}$, which implies (5.4).

6.4. Generic ranks. *Proof.* [Proof of theorem 5.9] Recall the morphism Σ_r defined by

$$\begin{aligned} \Sigma_r : (\widehat{X}_{(d-1,d)})^{\times r} &\rightarrow S^{d-1}V \oplus S^dV \\ ((\mu_1 \mathbf{v}_1^{d-1}, \lambda_1 \mathbf{v}_1^d), \dots, (\mu_r \mathbf{v}_r^{d-1}, \lambda_r \mathbf{v}_r^d)) &\mapsto (\mu_1 \mathbf{v}_1^{d-1} + \dots + \mu_r \mathbf{v}_r^{d-1}, \lambda_1 \mathbf{v}_1^d + \dots + \lambda_r \mathbf{v}_r^d). \end{aligned}$$

Let $\pi_{d-1} : S^{d-1}V \oplus S^dV \rightarrow S^{d-1}V$ be the natural projection, and likewise for π_d . Then $r \geq r_{\text{gen}}(X^{(d-1,d)})$ if and only if

$$\dim \pi_{d-1}(\pi_d^{-1}(p) \cap \text{Im } \Sigma_r) = \dim S^{d-1}V$$

for a general $p \in S^dV$. Since

$$\dim \pi_{d-1}(\pi_d^{-1}(p) \cap \text{Im } \Sigma_r) = \dim \text{Im}(\Sigma_r) - \dim S^dV \leq \dim(\widehat{X}_{(d-1,d)})^{\times r} - \dim S^dV,$$

then $r \geq \frac{\binom{m+d-2}{d-1} + \binom{m+d-1}{d}}{m+1}$. On the other hand,

$$\text{rank } p = \rho \stackrel{\text{def}}{=} r_{\text{gen}}(\nu_d(\mathbb{P}V)) = \lceil \frac{\binom{m+d-1}{d}}{m} \rceil,$$

so we may assume $p = \mathbf{u}_1^d + \dots + \mathbf{u}_\rho^d$ is a rank- ρ decomposition of p . Then inside $\pi_d^{-1}(p) \cap \text{Im } \Sigma_r$ there is a quasi-affine variety Y parametrized by

$$\begin{aligned} Y \stackrel{\text{def}}{=} \{ &(\mu_1 \cdot \mathbf{u}_1^{d-1} + \dots + \mu_\rho \cdot \mathbf{u}_\rho^{d-1} + \mathbf{u}_{\rho+1}^{d-1} + \dots + \mathbf{u}_r^{d-1}, \\ &\mathbf{u}_1^d + \dots + \mathbf{u}_\rho^d + 0 \cdot \mathbf{u}_{\rho+1}^d + \dots + 0 \cdot \mathbf{u}_r^d) \in S^{d-1}V \oplus S^dV : \\ &\mu_1, \dots, \mu_\rho \in \mathbb{C}, \mathbf{u}_{\rho+1}, \dots, \mathbf{u}_r \in V \}. \end{aligned}$$

Since $\dim Y \leq \dim \pi_d^{-1}(p) \cap \text{Im } \Sigma_r$, $\dim \pi_{d-1}(Y) \leq \dim \pi_{d-1}(\pi_d^{-1}(p) \cap \text{Im } \Sigma_r) \leq \dim S^{d-1}V$. Since $\rho \leq \dim S^{d-1}V$, p being general guarantees $\mathbf{u}_1^{d-1}, \dots, \mathbf{u}_\rho^{d-1}$ are linearly independent. Then when $r < \dim S^{d-1}V$ we can choose $\mathbf{u}_{\rho+1}, \dots, \mathbf{u}_r$ such that $\mathbf{u}_1^{d-1}, \dots, \mathbf{u}_r^{d-1}$ are linearly independent. By semicontinuity, for general $\mathbf{u}_{\rho+1}, \dots, \mathbf{u}_r$, we have $\mathbf{u}_1^{d-1}, \dots, \mathbf{u}_r^{d-1}$ are linearly independent. By Alexander-Hirschowitz theorem [1], when $r - \rho < r_{\text{gen}}(\nu_{d-1}(\mathbb{P}V))$, the quasi-affine variety parametrized by

$$\{\mathbf{u}_{\rho+1}^{d-1} + \dots + \mathbf{u}_r^{d-1} : \mathbf{u}_{\rho+1}, \dots, \mathbf{u}_r \in V\},$$

which contains an open Zariski subset of $\widehat{\sigma}_{r-\rho}(\nu_d(\mathbb{P}V))$, has the expected dimension $(r - \rho)m$. Therefore

$$\pi_{d-1}(Y) = \{\mu_1 \cdot \mathbf{u}_1^{d-1} + \dots + \mu_\rho \cdot \mathbf{u}_\rho^{d-1} + \mathbf{u}_{\rho+1}^{d-1} + \dots + \mathbf{u}_r^{d-1} : \mu_1, \dots, \mu_\rho \in \mathbb{C}, \mathbf{u}_{\rho+1}, \dots, \mathbf{u}_r \in V\}$$

has dimension $\rho + (r - \rho)m \leq \dim S^{d-1}V$, which implies

$$r_{\text{gen}}(X_{(d-1,d)}) \leq \left\lceil \frac{\binom{m+d-2}{d-1} + (m-1) \left\lceil \frac{\binom{m+d-1}{d}}{m} \right\rceil}{m} \right\rceil.$$

■

Proof. [Proof of theorem 5.10] Consider the isomorphism:

$$S^d(V \oplus \mathbb{C}) \cong S^{(0,\dots,d)}V.$$

Then we have that $\nu_d(V \oplus \mathbb{C})$ is isomorphic to

$$\widehat{Z}_d = \widehat{Z}_{d,V} \stackrel{\text{def}}{=} \{(c^d, c^{d-1}\mathbf{v}, \dots, c\mathbf{v}^{d-1}, \mathbf{v}^d) : c \in \mathbb{C}, \mathbf{v} \in V\} \subset \widehat{X}_{(0,\dots,d),V}. \quad (6.5)$$

Thus when $r > \dim S^{(0,\dots,d-2)}V$, for any $\mathbf{p} \in S^{(0,\dots,d-2)}V$ and any general

$$T = (c_1^d, \dots, \mathbf{v}_1^d) + \dots + (c_r^d, \dots, \mathbf{v}_r^d) \in \sigma_r(\widehat{Z}_d),$$

there are some $\alpha_1, \dots, \alpha_r$ such that

$$\mathbf{p} = \alpha_1(c_1^d, \dots, c_1^{d-2}\mathbf{v}_1^{d-2}) + \dots + \alpha_r(c_r^d, \dots, c_r^{d-2}\mathbf{v}_r^{d-2}) \in \sigma_r(\widehat{Z}_{d-2}) = S^{(0,\dots,d-2)}V,$$

which implies

$$\sigma_r(\widehat{Z}_d) = S^{(0,\dots,d-2)}V \oplus \sigma_r(\widehat{X}_{(d-1,d)}).$$

By Alexander-Hirschowitz Theorem, when $r < \frac{\binom{m+d}{d}}{m+1}$, $\dim \sigma_r(\widehat{Z}_d) = r(m+1)$. Therefore

$$\dim \sigma_r(\widehat{X}_{(d-1,d)}) = r(m+1) - \dim S^{(0,\dots,d-2)}V.$$

Since $m > (d-1)^2$, $\left\lceil \frac{\binom{m+d-2}{d-1} + \binom{m+d-1}{d}}{m+1} \right\rceil \geq \dim S^{(0,\dots,d-2)}V$. In particular,

$$r_{\text{gen}}(\widehat{X}_{(d-1,d)}) = \left\lceil \frac{\binom{m+d-2}{d-1} + \binom{m+d-1}{d}}{m+1} \right\rceil.$$

■

Proof. [Proof of theorem 5.11] When $r < \frac{\binom{m+d}{d}}{m+1}$, \widehat{Z}_d is r -identifiable, which implies when $r < r_{\text{gen}}(\widehat{X}_{(d-1,d)})$, $\widehat{X}_{(d-1,d)}$ is r -identifiable. ■

Acknowledgement. We would like to thank Ignat Domanov, Philippe Dreesen, Mariya Ishteva, Giorgio Ottaviani and Nick Vannieuwenhoven for enlightening discussions.

Appendix A. Basic definitions.

A.1. Symmetric tensors and homogeneous polynomials. Here we recall basic properties of symmetric tensors, that can be found in [16]. A tensor $\mathcal{T} \in \mathbb{K}^{m \times \dots \times m}$ is called symmetric if

$$\mathcal{T}_{i_1, \dots, i_d} = \mathcal{T}_{\pi(i_1, \dots, i_d)},$$

for any permutation of indices π . In this case, we write $\mathcal{T} \in S^d(\mathbb{K}^m)$. There is one-to one correspondence between symmetric tensors and homogeneous polynomials. The contraction

$$f(u) = \mathcal{T} \bullet_1 \mathbf{u} \bullet_2 \mathbf{u} \cdots \bullet_d \mathbf{u}.$$

is a homogenous polynomial of degree d . Vice versa, any homogeneous polynomial correspond to a unique element in $S^d(\mathbb{K}^m)$. In the paper, in order to avoid unnecessary extra symbols, for a homogeneous polynomial $f(\mathbf{u})$ we use the same letter for the corresponding $f \in S^d(\mathbb{K}^m)$.

Next, a rank-one symmetric tensor corresponds to the power of linear form:

$$(\mathbf{v} \otimes \dots \otimes \mathbf{v}) \bullet_1 \mathbf{u} \bullet_2 \mathbf{u} \cdots \bullet_d \mathbf{u} = (\mathbf{v}^\top \mathbf{u})^d,$$

which explains the equivalence between (1.4) and (1.3).

Finally, it is often convenient when the homogeneous polynomials are given in the following coordinates. Let $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$ be a multi index¹⁵, we define the set

$$\Delta_{s,m} \stackrel{\text{def}}{=} \{\alpha \in \mathbb{N}^m : \alpha_1 + \dots + \alpha_m = s\}$$

Now the homogeneous polynomial $f \in S^d V$ can be represented in the following coordinates

$$f(\mathbf{u}) = \sum_{\alpha=(\alpha_1, \dots, \alpha_m) \in \Delta_{d,m}} \frac{(\alpha_1 + \dots + \alpha_m)!}{\alpha_1! \cdots \alpha_m!} f_\alpha \mathbf{u}^\alpha,$$

where $\mathbf{u}^\alpha = u_1^{\alpha_1} \cdots u_d^{\alpha_d}$.

A.2. Algebraic varieties. This subsection gives an short summary of the definitions in algebraic geometry that will be needed in this paper. We choose a simplistic view, on a level of the popular book of Cox, Little and O'Shea [18]. A quick and simple overview of the main definitions used here can be also found in the paper [37]. As it was mentioned in the introduction, we only consider the case $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

Definition A.1 (Algebraic variety). A subset $Z \subseteq \mathbb{K}^N$ is called an affine algebraic variety¹⁶ if there exist a finite set of polynomials $p_1, \dots, p_M \in \mathbb{K}[\mathbf{z}]$ such that

$$\mathbf{z} \in Z \iff \begin{cases} p_1(\mathbf{z}) &= 0, \\ &\vdots \\ p_M(\mathbf{z}) &= 0 \end{cases} \quad (\text{A.1})$$

¹⁵By convention, the set \mathbb{N} includes 0.

¹⁶As in [18], we do not require a variety to be irreducible, contrary to some classic definitions.

i.e. Z is a zero locus of p_1, \dots, p_M . A set X is called a proper subvariety of Z , if $X \subsetneq Z$ and X is also a variety.

Remark 15. \mathbb{K}^N is also an algebraic variety: a zero locus of the zero polynomial $p(\mathbf{z}) \equiv 0$.

Definition A.2 (Zariski closure). For any set $Y \in \mathbb{K}^N$, by \bar{Y} we denote the smallest by inclusion algebraic variety Z , such that $Y \subseteq Z$. \bar{Y} is called the Zariski closure of Y .

Definition A.3 (Irreducibility). A nonempty variety X is called irreducible [18] if it cannot be represented as a union of two distinct varieties. (More precisely, if for a decomposition $X = Y \cup Z$ with Y, Z varieties, it holds that either $Y \subseteq Z$ or $Z \subseteq Y$.)

Definition A.4 (Generic property). We say that some property is generic in an irreducible variety Z if there exists a proper subvariety $V \subsetneq Z$ (of smaller dimension) such that the property is true for all points in $Z \setminus V$.

Remark 16 (Generic properties in \mathbb{C}^N). If the property is generic in \mathbb{C}^N , it implies¹⁷ that a random vector in \mathbb{C}^N (drawn from any absolutely continuous distribution) satisfies a given generic property with probability 1.

Definition A.5 (complexification). For a real variety $X_{\mathbb{R}} \subset \mathbb{R}^N$ (given by real polynomials in (A.1)) its complexification $X_{\mathbb{C}} = X_{\mathbb{R}} \otimes \mathbb{C}$ is the variety in \mathbb{C}^N cut out by the same system of equations.

Definition A.6. Let Z be an irreducible variety in \mathbb{K}^N , and p_1, \dots, p_M are the generators of its ideal. Let d be the maximal rank of the Jacobian matrix $J_p(\mathbf{z}) \stackrel{\text{def}}{=} [\frac{\partial p_j}{\partial z_i}]_{i,j=1}^{M,N}$ at $\mathbf{z} \in Z$. Then the dimension is, by definition, $N - d$. The dimension of a reducible variety is equal to the maximal dimension of its irreducible components.

A.3. Polynomial images of algebraic varieties. **Definition A.7.** The set $Z \in \mathbb{C}^N$ is called constructible, if it can be written as a finite union

$$Z = (X_1 \setminus Y_1) \cup \dots \cup (X_r \setminus Y_r)$$

where X_k, Y_k are varieties.

Theorem A.8 (Chevalley). The polynomial image of a constructible set is constructible.

Corollary A.9. Assume that $X \subset A$, $p : A \rightarrow B$ is a polynomial map and Y is the Zariski closure $Y = \overline{p(X)}$, such that Y is irreducible. Then a general element in Y lies in $p(X)$, i.e. there exists a subvariety $Z \subsetneq Y$ of strictly smaller dimension such that $Y \setminus Z \in p(X)$.

REFERENCES

- [1] JAMES ALEXANDER AND ANDRÉ HIRSCHOWITZ, *Polynomial interpolation in several variables*, Journal of Algebraic Geometry, 4 (1995), pp. 201–222.
- [2] ELENA ANGELINI, FRANCESCO GALUPPI, MASSIMILIANO MELLA, AND GIORGIO OTTAVIANI, *On the number of Waring decompositions for a generic polynomial vector*, tech. report, arxiv.org, 2016. Available from <http://arxiv.org/abs/1601.01869>.
- [3] EDOARDO BALlico, *On the weak non-defectivity of veronese embeddings of projective spaces*, Central European Journal of Mathematics, 3 (2005), pp. 183–187.
- [4] ALESSANDRA BERNARDI, GRIGORIY BLEKHERMAN, AND GIORGIO OTTAVIANI, *On real typical ranks*, tech. report, arxiv.org, 2015. Available from <http://arxiv.org/abs/1601.01869>.

¹⁷This follows from the fact that any proper algebraic subvariety has Lebesgue measure zero.

- [5] ANDRZEJ BIAŁYŃICKI-BIRULA AND ANDRZEJ SCHINZEL, *Representations of multivariate polynomials as sums of polynomials in linear forms*, Colloq. Mathematicum, 112 (2008), pp. 201–233.
- [6] GRIGORIY BLEKHERMAN AND ZACH TEITLER, *On maximum, typical and generic ranks*, Mathematische Annalen, 362 (2015), pp. 1021–1031.
- [7] CRISTIANO BOCCI, LUCA CHIANTINI, AND GIORGIO OTTAVIANI, *Refined methods for the identifiability of tensors*, Annali di Matematica Pura ed Applicata (1923–), 193 (2014), pp. 1691–1702.
- [8] ENRICO CARLINI, MARIA VIRGINIA CATALISANO, AND ANTHONY V. GERAMITA, *The solution to the waring problem for monomials and the sum of coprime monomials*, Journal of Algebra, 370 (2012), pp. 5 – 14.
- [9] MICHAEL L. CATALANO-JOHNSON, *The possible dimensions of the higher secant varieties*, American Journal of Mathematics, (1996), pp. 355–361.
- [10] SCOTT SHAOBING CHEN, DAVID L. DONOHO, AND MICHAEL A. SAUNDERS, *Atomic decomposition by basis pursuit*, SIAM Review, 43 (2001), pp. 129–159.
- [11] LUCA CHIANTINI AND CIRO CILIBERTO, *Weakly defective varieties*, Transactions of the American Mathematical Society, 354 (2002), pp. 151–178.
- [12] LUCA CHIANTINI AND CIRO CILIBERTO, *On the concept of k -secant order of a variety*, Journal of the London Mathematical Society, 73 (2006), pp. 436–454.
- [13] LUCA CHIANTINI, GIORGIO OTTAVIANI, AND NICK VANNIEUWENHOVEN, *An algorithm for generic and low-rank specific identifiability of complex tensors*, SIAM Journal on Matrix Analysis and Applications, 35 (2014), pp. 1265–1287.
- [14] LUCA CHIANTINI, GIORGIO OTTAVIANI, AND NICK VANNIEUWENHOVEN, *On generic identifiability of symmetric tensors of subgeneric rank*, Transactions of the American Mathematical Society, (2016). to appear.
- [15] PIERRE COMON, JOS M. F. TEN BERGE, LIEVEN DELATHAUWER, AND JOSEPHINE CASTAING, *Generic and typical ranks of multi-way arrays*, Linear Algebra Appl., 430 (2009), pp. 2997–3007.
- [16] PIERRE COMON, GENE H. GOLUB, LEK-HENG LIM, AND BERNARD MOURRAIN, *Symmetric tensors and symmetric tensor rank*, SIAM. J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279.
- [17] PIERRE COMON, YANG QI, AND KONSTANTIN USEVICH, *A polynomial formulation for joint decomposition of symmetric tensors of different orders*, in Latent Variable Analysis and Signal Separation, Emmanuel Vincent, Arie Yeredor, Zbyněk Koldovský, and Petr Tichavský, eds., vol. 9237 of Lecture Notes in Computer Science, Springer, 2015, pp. 22–30.
- [18] DAVID COX, JOHN LITTLE, AND DONALD O’SHEA, *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer, 2nd ed., 1997.
- [19] LIEVEN DE LATHAUWER, *A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization*, SIAM Journal on Matrix Analysis and Applications, 28 (2006), pp. 642–666.
- [20] PHILIPPE DREESEN, MARIYA ISHTEVA, AND JOHAN SCHOUKENS, *Decoupling multivariate polynomials using first-order information*, SIAM. J. Matrix Anal. Appl., 36 (2015), pp. 864–879.
- [21] RICHARD EHRENBORG AND GIAN-CARLO ROTA, *Apolarity and canonical forms for homogeneous polynomials*, European Jour. Combinatorics, 14 (1993), pp. 157–181.
- [22] DAVID EISENBUD, *Linear sections of determinantal varieties*, American Journal of Mathematics, 110 (1988), pp. 541–575.
- [23] WILLIAM FULTON AND JOE HARRIS, *Representation theory: a first course*, Springer Science & Business Media, 2013.
- [24] F. GIRI AND E.W. BAI, *Block-oriented Nonlinear System Identification*, Lecture Notes in Control and Information Sciences, Springer, 2010.
- [25] ROBIN HARTSHORNE, *Algebraic geometry*, Springer-Verlag, New York-Heidelberg, 1977. Graduate Texts in Mathematics, No. 52.
- [26] ANTHONY IARROBINO AND VASSIL KANEV, *Power sums, Gorenstein Algebras and Determinantal Loci*, vol. 1721 of Lecture Notes in Mathematics, Springer, 1999.
- [27] JOSEPH M. LANDSBERG, *Tensors: Geometry and applications*, vol. 128, American Mathematical Soc., 2012.
- [28] JOSEPH M. LANDSBERG AND LAURENT MANIVEL, *On the ideals of secant varieties of Segre varieties*, Foundations of Computational Mathematics, 4 (2004), pp. 397–422.

- [29] BENJAMIN F. LOGAN AND LARRY A. SHEPP, *Optimal reconstruction of a function from its projections*, Duke Math. J., 42 (1975), pp. 645–659.
- [30] MASSIMILIANO MELLA, *Singularities of linear systems and the waring problem*, Transactions of the American Mathematical Society, 358 (2006), pp. 5523–5538.
- [31] KONSTANTIN I. OSKOLKOV, *On representations of algebraic polynomials as a sum of plane waves*, Serdica Mathematical Journal, (2002), pp. 379–390.
- [32] YANG QI, PIERRE COMON, AND LEK-HENG LIM, *Semialgebraic geometry of nonnegative tensor rank*, SIAM Journal on Matrix Analysis and Applications, 37 (2016), pp. 1556–1580.
- [33] ANDRZEJ SCHINZEL, *On a decomposition of polynomials in several variables*, Journal de Théorie de Nombres de Bordeaux, 14 (2002), pp. 647–666.
- [34] ———, *On a decomposition of polynomials in several variables, ii*, Colloquium Mathematicum, 92 (2002), pp. 67–79.
- [35] JOHAN SCHOUKENS, ANNA MARCONATO, RIK PINTELON, ET AL., *System identification in a real world*, in IEEE 13th International Workshop on Advanced Motion Control (AMC), March 2014, pp. 1–9.
- [36] YOAN SHIN AND JOYDEEP GHOSH, *Ridge polynomial networks*, IEEE Transactions on Neural Networks, 6 (1995), pp. 610–622.
- [37] FRANK SOTTILE, *Real algebraic geometry for geometric constraints*, tech. report, 2016. arXiv preprint 1606.03127.
- [38] V. STRASSEN, *Rank and optimal computation of generic tensors*, Linear Algebra and its Applications, 5253 (1983), pp. 645 – 685.
- [39] KONSTANTIN USEVICH, *Decomposing multivariate polynomials with structured low-rank matrix completion*, in 21st Int. Symposium on Mathematical Theory of Networks and Systems, July 7–11, 2014. Groningen, The Netherlands, 2014, pp. 1826–1833.
- [40] ANNE VAN MULDER, LAURENT VANBEYLEN, AND KONSTANTIN USEVICH, *Identification of a block-structured model with several sources of nonlinearity*, in Proceedings of the 14th European Control Conference (ECC 2014), 2014, pp. 1717–1722.
- [41] FYODOR L. ZAK, *Determinants of projective varieties and their degrees*, in Algebraic Transformation Groups and Algebraic Varieties, V. L. Popov, ed., Springer, Berlin, 2004, pp. 207–238.